

# Flavio Vinicius Diniz de Figueiredo

**e-mail:** flaviovdf [AT] gmail.com    **web:** <http://flaviovdf.github.io>

---

- EDUCATION**
- PhD*, Computer Science 2010 to 2015  
Universidade Federal de Minas Gerais (UFMG)  
Belo Horizonte, MG, Brasil  
Advisor: Prof. Jussara Marques de Almeida  
GPA: 88.2 out of 100<sup>†</sup>
- MsC*, Computer Science 2008 to 2010  
Universidade Federal de Minas Gerais (UFMG)  
Belo Horizonte, MG, Brasil  
Advisor: Prof. Jussara Marques de Almeida  
GPA: 88.7 out of 100
- BsC*, Computer Science 2003 to 2007  
Universidade Federal de Campina Grande (UFCG)  
Campina Grande, PB, Brasil  
Advisor: Prof. Francisco Vilar Brasileiro  
GPA: 76.3 out of 100
- LANGUAGES**    Fluent in English and Portuguese.
- CURRENT POSITION**    **Professor of Computer Science at Federal University of Minas Gerais (UFMG)**  
Nov 2016 - Now
- Professor of Computer Science starting January 2016. Will teach and research on computer science and data science courses.
  - Keywords: Computer Science, Data Science, Machine Learning
- PAST R & D EXPERIENCE**
- Research Staff Member at IBM Research — Brasil**    Nov 2015 - Nov 2016
- Member of the social data analytics team at IBM Research Rio de Janeiro. Experimented with and implemented novel information retrieval techniques used to support chatbot systems. Also studied the efficacy of social media data as supporting sources of information for multi agent chat bots.
  - Two patents filed
  - Keywords: Machine Learning, Data Science, Information Retrieval, Chatbots, Java, Golang, Python, Numpy's data stack
- PostDoc Researcher at UFCG**    June 2015 to November 2015
- Postdoctoral researcher. Supervised PhD and MsC students in data science and computational urban studies. Also supervised students in an initial prototype for a academic recommender system for students.
  - Keywords: Machine Learning, Data Science, Recommender systems
- PhD Researcher at UFMG**    July 2010 to May 2015
- PhD Thesis on mining social media popularity. Implemented novel techniques for mining large datasets of time series extracted from popular websites at the time. Published papers in top academic conferences and journals.

---

<sup>†</sup>GPA's on a 0 to 100 scale. Not graded on a curve.

- Research done for Brazil’s National Institute for Web Research (InWeb - <http://www.inweb.org.br>)
- Keywords: Machine Learning, Data Science, Data Mining, Web Crawling, Java, Python
- Open source code:  
<http://github.com/flavioovdf/spiderpig>  
<http://github.com/flavioovdf/youtime>  
<http://github.com/flavioovdf/pyksc>  
<http://github.com/flavioovdf/phoenix>  
<http://github.com/flavioovdf/tribeflow>  
<http://github.com/flavioovdf/socialmediapop>

**Visiting Scholar at Carnegie Mellon University** October 2013 to October 2014

- Visiting the Databases Group (<http://db.cs.cmu.edu>)
- Researching social media access time series. Creating models for time series evolution and prediction.
- Joint work done with Prof. Christos Faloutsos
- Keywords: Data mining, Time series, Machine Learning
- Open sourced code:  
<http://github.com/flavioovdf/phoenix>

**Visiting Scholar at University of British Columbia** March to August 2011

- Visiting the Networked Systems Laboratory (<http://netsyslab.ece.ubc.ca>)
- Researching tagging systems (i.e. Delicious) from a online peer production systems perspective. In details, we sought to measure the value of information in tags for peers (users) in the system
- Joint work done with Prof. Matei Ripeanu and PhD student Elizeu Santos-Neto
- Keywords: Information Retrieval, NLP, Python, Cython
- Open sourced code:  
[http://github.com/flavioovdf/tag\\_assess](http://github.com/flavioovdf/tag_assess)

**MsC Researcher at UFMG** 2008 to Feb 2010

- Research done on social network and Web 2.0 sites with the main objective of extracting evidence of quality of textual data (i.e. tags or comments) data when applied to information retrieval (IR) tasks.
- Research done for Brazil’s National Institute for Web Research (InWeb - <http://www.inweb.org.br>)
- Keywords: Data Mining, Web Crawling, Information Retrieval, Java, Python
- Open sourced code:  
<http://github.com/flavioovdf/spiderpig>

**R & D for OurGrid** 2007

- Member the OurGrid Project (<http://www.ourgrid.org>)
- Researcher on cloud and utility computing. More specifically, we studied how to create hybrid clouds using local, shared and utility computing resources.
- Keywords: Distributed Systems, Cloud Computing, Virtual Machines, Java

- Open sourced code:  
<http://github.com/ourgrid/ourgrid>

#### **Developer for OurGrid**

2006

- Member the OurGrid Project (<http://www.ourgrid.org>)
- Part of the development team which at the time was focused on re-coding the the OurGrid middleware for a new event based communication infrastructure.
- Keywords: Distributed Systems, Java, RMI
- Open sourced code:  
<http://github.com/ourgrid/ourgrid>

#### **Developer at Failure Spotter**

2005

- Member the Failure Spotter Project (<http://spotter.lsd.ufcg.edu.br>)
- Developed a fault tolerant execution and scheduling service for the grid middleware OurGrid (<http://www.ourgrid.org>).
- Keywords: Distributed Systems, Fault Tolerance, Java, RMI
- Open sourced code:  
<http://github.com/ourgrid/ourgrid>

#### **TECHNICAL SKILLS**

**Programming Languages** Golang, Java, Python, Unix Scripting (i.e. Bash)

**Data Analysis** Numpy's data stack, Octave and R

**Areas of Expertise** Data Science, Data Mining, Machine Learning, Parallel Computing, Distributed computing, Web crawling, Information Retrieval (IR)

#### **OPEN SOURCE CODE**

Some repositories were listed above in their respective project. For a full list please visit: <http://github.com/flaviovdv>

#### **AWARDS**

**2016** Best Paper Award at the ACM WebScience Conference

**2015** Best thesis in Computer Science at the Federal University of Minas Gerais

**2014** Honorable Mention as one of the best paper at the ACM Digital Library Conference

**2014** Google Latin America Ph.D. Fellowship Award – Renewal

**2014** First Position at the ECML/PKDD Predictive Analytics Challenge

**2013** Capes Scholarship for Visiting Scholar (PhD. Sandwich Program)

**2013** Honorable Mention at the First Open Data Science Competition from the Brazilian Ministry of Justice

**2013** Google Latin America Ph.D. Fellowship Award

**2012** IBM Ph.D. Fellowship Award

**2011** Canadian Emerging Leaders in the Americas Program (ELAP)

**2011** CNPq Scholarship for PhD Studies

**2010** Capes Scholarship for PhD Studies

**2010** National Institute for Web Research (InWeb) Scholarship for MSc studies

**2009** UOL Scholarship for MSc Studies. Grant 20090215103600

**2008** CNPq Scholarship for MSc Studies

**2006 to 2007** Hewlett Packard Brazil Scholarship for the OurGrid project

**2005** Hewlett Packard Brazil scholarship for the Failure Spotter project

**JOURNAL  
PUBLICATIONS**

Figueiredo, F., Almeida, J., Benevenuto, F., Gonçalves, M. “TrendLearner: Early Prediction of Popularity Trends of User Generated Content”, *Elsevier Information Sciences*, Volumes 349–350

Figueiredo, F., Almeida, J., Benevenuto, F., Gonçalves, M. “On the Dynamics of Social Media Popularity: A YouTube Case Study”, *ACM Transaction on Internet Technology*, Vol. 14, Issue 4, December 2014.

Figueiredo, F., Belém, F., Pinto, H., Almeida, J., Gonçalves, M., Fernandes, D. and Moura, E. “Assessing the Quality of Textual Features in Social Media”, *Information Processing & Management*, Vol. 49, pp. 222 - 247, 2013.

Almeida, J., Gonçalves, M., Figueiredo, F., Belém, F. and Pinto, H. “On the Quality of Information for Web 2.0 Services”, *IEEE Internet Computing*, Vol. 14, pp. 47-55, 2010.

**SELECTED  
CONFERENCE  
PUBLICATIONS**

For a complete list visit: <http://www.dcc.ufmg.br/~flaviov>

Figueiredo, F., Ribeiro, B., Almeida, J., Andrade, N., Faloutsos, C. “Mining Online Music Listening Trajectories” *International Society for Music Information Retrieval - ISMIR 2016*.

Figueiredo, F., Andrade, N. “Exploring the Latent Structure of Collaborations in Music Recordings: A Case Study in Jazz” *International Society for Music Information Retrieval - ISMIR 2016*.

Figueiredo, F., Ribeiro, B., Almeida, J., Faloutsos, C. “TribeFlow: Mining & Predicting User Trajectories” *ACM World Wide Web Conference - WWW 2016*

Figueiredo, F., Ribeiro, B., Almeida, J., Faloutsos, C. “A Summary of the TribeFlow Model for Music Discovery Applications” *Machine Learning for Music Discovery Workshop @ ICML 2016*

Arantes, M., Figueiredo, F., Almeida, J. “Understanding Video-Ad Consumption on YouTube: A Measurement Study on User Behavior, Popularity, and Content Properties.” *ACM Web Science Conference - WebSci 2016*

Cavalin, P., Figueiredo, F., Bayser, M., Moyano, L., Candello, H., Appel, A., Souza, R. “Building a Question-Answering Corpus using Social Media and News Articles” *International Conference on the Computational Processing of Portuguese Language - PROPOR 2016*

Figueiredo, F., Almeida, J., Gonçalves, M. “Improving the Effectiveness of Content Popularity Prediction Methods using Time Series Trends “ In *ECML/PKDD Predictive Analytics Challenge*, 2014.

Figueiredo, F., Matsubara, Y., Ribeiro, B., Almeida, J., Faloutsos, C. “Revisit Behavior in Social Media: The Phoenix-R Model and Discoveries “ In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery -*

*ECML/PKDD*, 2014.

Gonçalves, G., Figueiredo, F., Almeida, J., Gonçalves, M. “Characterizing Scholar Popularity: A Case Study in the Computer Science Research Community” In *ACM/IEEE Joint Conference on Digital Libraries - JCDL*, 2014.

Figueiredo, F., Almeida, J., Benevenuto, F., Gummadi, K. “Does Content Determine Information Popularity in Social Media? A Case Study of YouTube Videos’ Content and their Popularity ” In *Proc. ACM CHI Conference on Human Factors in Computing Systems - CHI*, 2014.

Figueiredo, F. “On the Prediction of Popularity of Trends and Hits for User Generated Videos” In *Proc. ACM Conference on Web Search and Data Mining - WSDM*, 2013.

Figueiredo, F., Benevenuto, F. and Almeida, J. “The Tube over Time: Characterizing Popularity Growth of YouTube Videos”, In *Proc. ACM Conference on Web Search and Data Mining - WSDM*, 2011.

Santos-Neto, E., Figueiredo, F., Almeida, J., Mowbray, M., Gonçalves, M. and Rippeanu, M. “Assessing the Value of Contributions in Tagging Systems”, In *Proc. IEEE International Symposium on Social Intelligence and Networking - SIN*, 2010.

Figueiredo, F., Belém, F., Pinto, H., Almeida, J., Gonçalves, M., Fernandes, D., Moura, E. and Cristo, M. “Evidence of Quality of Textual Features on the Web 2.0”, In *Proc. ACM Conference on Information Knowledge and Management - CIKM*, 2009.

Ditarso, P., Figueiredo, F., Candeia, D., Brasileiro, F. and Coelho, A. “On the Planning of a Hybrid IT Infrastructure”, In *Proc. IEEE Network Operations and Management Symposium - NOMS*, 2008.

Coelho, A., Ditarso, P., Figueiredo, F., Candeia, D. and Brasileiro, F. “On the Impact of Choice in Multi-Service P2P Grids”, In *Proc. Workshop on Business Driven IT Management in Conjunction with NOMS*, 2008.

Brasileiro, F., Araújo, E., Voorsluys, W., Oliveira, M. and Figueiredo, F. “Bridging the High Performance Computing Gap: the OurGrid Experience”, In *Proc. First Latin America Grid Workshop - LAGrid07 with CCGRID*, 2007.

Figueiredo, F., Brito, A. and Brasileiro, F. “Implementing a Distributed Execution Service for a Grid Broker”, In *Proc. Fault Tolerance Workshop in conjunction with the Brazilian Symposium of Computer Networks*, 2006.

**REFERENCES**      Provided upon request